

The function of the self-attention network

Sheila J. Cunningham

This is an Accepted Manuscript of an article published by Taylor & Francis in Cognitive Neuroscience on 17th August 2015, available online:

<http://www.tandfonline.com/10.1080/17588928.2015.1075485>

RUNNING HEAD: The Function of the SAN

The Function of the Self-Attention Network

Sheila J. Cunningham

Psychology Division, Abertay University

COMMENTARY

On Humphreys & Sui's 'Attentional control and the self: The Self-Attention Network (SAN)

(710 words, including references)

Correspondence address:

Division of Psychology,

Abertay University,

Bell Street, Dundee,

DD1 1HG

U.K.

Email: s.cunningham@abertay.ac.uk

ABSTRACT

This commentary links Humphrey and Sui's proposed Self-attention Network (SAN) to the memory advantage associated with self-relevant information (i.e., the self-reference effect). Articulating this link elucidates the functional quality of the SAN in ensuring that information of potential importance to self is not lost. This adaptive system for self-processing mirrors the cognitive response to threat stimuli, which also elicit attentional biases and produce characteristically enhanced, episodic representations in memory. Understanding the link between the SAN and memory is key to comprehending more broadly the operation of the self in cognition.

The Function of the Self-Attention Network

Our understanding of the cognitive systems triggered by perception of self-relevant cues has grown significantly in the past 10 years, as Humphreys and Sui's timely target article makes clear. Their proposed 'Self-Attention Network' (SAN) is built on a solid foundation of recent behavioral and neural research on self-attention biases, which reveals a combination of top-down and bottom-up processes involving both the attention control network and areas associated with self-referential cognition. While the SAN framework is a valuable step forward in refining our understanding of self-processing biases, one important consideration that is not addressed in the target article is the *function* the SAN, an aspect that should be central to any big-picture conceptualization of self-processing biases.

I submit that the SAN functions to ensure that information of potential relevance to self is preferentially encoded; in other words, attention biases serve to support the self-reference effect in memory (Symons & Johnston, 1997). This robust memory advantage for information associated with self at encoding is elicited by even minimal self-stimulus connections, such as simultaneous presentation of stimuli with a self-cue, or temporary ownership in a sorting game (Cunningham, Turk, MacDonald, & Macrae, 2008; Turk, Cunningham, & Macrae, 2008). It serves the important ecological function of ensuring that information of potential use to self is not lost (Cunningham, Brady-Van den Bos, Gill, & Turk, 2013).

This functional explanation of the SAN mirrors other attention-based memory effects, such as the cognitive response to threat stimuli. When a threatening cue is perceived it captures

attention even when the cue is not task-relevant, because survival is a constant underlying goal (Öhman & Mineka, 2001). Increases in attentional focus and neurotransmitter levels in response to threat cues result in memories that are characterized by binding and episodic recollection (see Hadley & MacKay, 2006). The effects of self on cognition strikingly echo this system: Self-relevance is perpetually goal-relevant, self-cues attract attention via the SAN and resultant memories are characteristically episodic in nature, bound with information from the encoding context (Conway & Dewhurst, 1995).

Supporting the idea that the self improves episodic binding in a manner akin to emotional processing, we have shown that while centrally presented self-cues (e.g., own-name, own face) attract attention, they do not detract from the processing of simultaneously presented peripheral stimuli as might be expected; rather, these are remembered better than stimuli presented with cues of other-referents (Turk et al., 2008). Interestingly, recent data (Cunningham & Allan, 2015) suggests that when stimuli are presented with self- and other-faces at short SOAs (< 250ms), the usual memory advantage for those presented with self over other does not emerge. However when the SOA is longer (700ms), memory for other-referent stimuli decreases significantly while self-referent stimuli remain well-remembered, creating a self-reference effect. This suggests that the self continues to bind the stimuli and cue together when the binding produced by temporal proximity has lapsed.

Extending Humphreys and Sui's argument, information relevant to self is subject to biases that combine to produce a robust and adaptive memory advantage. Understanding the link

between the SAN and memory is key to comprehending more broadly the operation of the self in cognition.

References

Conway, M. A. & Dewhurst, S. A. (1995). The self and recollective experience. *Applied Cognitive Psychology*, 9, 1-19. doi:10.1002/acp.2350090102.

Cunningham, S. J., & Allan, K. (2015). *Temporal proximity and the incidental self-reference effect*. Unpublished manuscript.

Cunningham, S. J., Brady-Van den Bos, M., Gill, L., & Turk, D. J. (2013). Survival of the selfish: Contrasting self-referential encoding and survival processing. *Consciousness and Cognition*, 22, 237-244. doi: 10.1016/j.concog.2012.12.005.

Cunningham, S. J. Turk, D. J., Macdonald, L. M. & Macrae, C. N. (2008). Yours or mine? Ownership and memory. *Consciousness and Cognition*, 17, 312-318. doi: 10.1016/j.concog.2007.04.003.

Hadley, C. B. & MacKay, D. G. (2006). Does emotion help or hinder immediate memory? Arousal versus priority-binding mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 79 – 88. doi: 10.1037/0278-7393.32.1.79.

Öhman, A. & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, 108, 483-522. doi: 10.1037/0033-295X.108.3.483.

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, 121, 371-394. doi:10.1037/0033-2909.121.3.371

Turk, D. J., Cunningham, S. J. & Macrae, C. N. (2008). Self-memory biases in explicit and incidental encoding of trait adjectives. *Consciousness and Cognition*, 17, 1040-1045.
doi: 10.1016/j.concog.2008.02.004.